Published online September 14, 2004

4876–4883 *Nucleic Acids Research, 2004, Vol. 32, No. 16*
doi:10.1093/nar/gkh826

DNA base pair resolution by single molecule force spectroscopy

Bernie D. Sattin, Andrew E. Pelling and M. Cynthia Goh*

Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada

Received June 1, 2004; Revised July 23, 2004; Accepted August 26, 2004

ABSTRACT

The forces that hold complementary strands of DNA together in a double helix, and the role of base mismatches in these, are examined by single molecule force spectroscopy using an atomic force microscope (AFM). These forces are important when considering the binding of proteins to DNA, since these proteins often mechanically stretch the DNA during their action. In AFM measurement of forces, there is an inherent instrumental limitation that makes it difficult to compare results from different experimental runs. This is circumvented by using an oligonucleotide microarray, which allowed a direct comparison of the forces between perfectly matched short oligonucleotides and those containing a single or double mismatch. Through this greatly increased sensitivity, the force contribution of a single AT base pair was derived. The results indicate that the contribution to forces from the stacking interactions is more important than that from hydrogen bonding.

INTRODUCTION

The intermolecular forces that hold DNA in a duplex are fundamental to life's processes. DNA not only encodes the machinery that makes life work, but also controls the exact proportion of each necessary ingredient. Ever since the defining works of Breslauer and others (1,2), the sequence-dependent thermodynamics of short DNA duplexes have been probed through melting temperatures and calorimetry [e.g. (3–5)]. There has been a recent resurgence of activity in this area with the use of force spectroscopy (FS) techniques, using either the atomic force microscope (AFM) or laser optical traps (optical tweezers), which fall into two general categories: those conducted on short DNA and those on long DNA. Experiments on long DNA focused on the overall properties of the molecule (6,7) and resulted in the discovery S-DNA (8–11). Those on short duplexes focused on the reaction pathway of melting and contributed to the understanding of the local unwinding of DNA (12–17). In this work, we push the limits of FS to derive the forces between strands of DNA due to a single base pair.

The pioneering work in the use of the AFM for FS of DNA was performed by Lee *et al*. (12). As one of the earliest FS experiments, there are issues that made these results difficult to relate to recent efforts. The experimental design, however, was clever: they used a repeating DNA sequence [$(ACTG)_5$] that could form duplexes of several unique lengths (12, 16, 20 bp), which would give rise to several unique binding forces. Thus, three different interactions could be measured and compared within the same experiment. The forces measured were 765, 655 and 415 pN, for helices of 20, 16 and 12 bp, respectively, under slightly ionic conditions (0.1 N NaCl) (13). In the experiments performed later (14–17) the expected value for the rupture of short duplex DNA (10–30 bp) was reduced down to the level of 20–200 pN, depending on the length and conditions. To estimate the magnitude of forces holding together single base pairs, experiments were conducted using AFM (18) on a self-assembled purine/pyrimidine monolayer-coated tip and surface, and by optical tweezers (11) or AFM (9) on long tracts of GC or AT base pairs. The former approach yielded a value of ~54 pN (18), while the latter gave ~9–20 pN (9,11).

An issue inherent to all AFM-based force measurements is that small differences in conditions and in the cantilever spring constant can drastically affect results (19), such that studies carried out using different probe tips or under different solution conditions cannot be compared quantitatively. This is particularly relevant when one is interested in very small forces, such as those between a single base pair. We will discuss the factors that are relevant to force measurement, and the experimental design we used to overcome these factors and enable us to obtain significant results.

To obtain forces by AFM, a value for the cantilever spring constant is needed, which has to be obtained by calibration (20–24). The expected error in these calibrations is ~5% at best, and typically up to 20% (19). This uncertainty will propagate when comparing measurements using different cantilevers. Furthermore, using different solutions and conditions, the spring constant may change and this change may or may not be negligible. Thus, changes in buffer concentration and aging can strongly influence comparability of different experimental runs. In addition to affecting the spring constant directly, non-specific forces due to charges (ions, surfaces) can influence the measured forces.

*To whom correspondence should be addressed. Tel: +1 416 978 6254; Email: cgoh@chem.utoronto.ca
Present address:
Andrew E. Pelling, UCLA, Department of Chemistry and Biochemistry, 607 Charles E. Young Dr East, Los Angeles, CA 90095, USA

Immobilization chemistry is important since the interacting species have to be attached to the surface and to the probe tip, and this can play a large role in the measurement. In some studies, oligodeoxynucleotides (ODNs) were covalently immobilized directly on glass, while in others, a long polyethyleneglycol spacer was used as linker. Glass supports a surface charge under typical conditions for these studies, and this can influence the measurements either directly, or by its interaction with the species under study. Having a linker moves the ODN away from the surface and gives it more freedom to interact. But the linker itself may change the nature of the interaction, or complicate data interpretation, as it has to be unwound or stretched before or at the same time as the molecule of interest. Another way of attachment is through a thiolated linker interacting with a gold-coated surface; in this case, the interactions of the metal layer could be important.

Finally, in measurements of very small forces, the collection of large data sets is important in order to derive proper statistics. This is sometimes difficult to accomplish because of the difficult and repetitive nature of the experiments and the technical difficulties that could arise during data collection.

Our goal is to obtain values for forces between single base pairs and in order to do so, the issues raised above have to be addressed. We designed an experimental approach that allowed us to do so.

The plan is to derive the role of single base pair forces by comparing perfectly matched ODN pairs to those with mismatches. The above analysis shows that the most significant problem to overcome is comparing the different experiments so as to get very small numbers, when the random errors that are generated due to cantilever calibration, solution, surface and sample conditions, etc., produce large uncertainties in the data. However, note that these are random errors that reduced significantly by using the same cantilever for all the measurements, and by conducting all the measurements under exactly the same conditions. This can be done if all the experiments are performed within the same solution. We, thus, built a small microarray of ODNs that fits within the AFM fluid cell, and used the same cantilever to measure forces between the ODN on the probe tip and on each of the microarray elements.

We followed the work of Lee *et al*. (12), and used the same tip and surface ODNs to measure multiple interactions and to enable comparisons to their work. We deviated from them by using gold/sulfur chemistry to immobilize ODNs, to reduce interference from non-specific tip/surface interactions (15). We used a hexacarbon linker to separate the ODNs from the influence of the gold surfaces and to give them flexibility for interaction. The short linker was chosen to reduce complications that could arise due to stretching and unwinding. We performed hundreds of repeated measurements and introduced cluster analysis for analyzing our data.

In this paper, we report the results of the improvements we introduced to FS that allowed us to make measurements of higher precision than all previous ones. This increased resolution enabled us to determine the rupture force of a single AT base pair.

MATERIALS AND METHODS

Materials

Single-stranded ODNs were from L.-C. Tsui (The Centre for Applied Genomics, Toronto), synthesized using standard procedures (25,26). The ODNs had a hexanethiol attached to the 3′ end and were received in a pure, dry form. They were dissolved in ultrapure water (18 mΩ) to 1 mg/ml for short-term storage at −20°C. Four different sequences were synthesized: TIP $(ACTG)_5$, TIP's complement MATCH $(CAGT)_5$, TIP's complement with a single mismatch near the center 1MIS (CAGTCAGTCAGCCAGTCAGT) and TIP's complement with a double mismatch near the center 2MIS (CAGTCAGTCAACCAGTCAGT). Stock solutions of ODN (1.6 nM) were made by diluting 10 and 20 µl of ODN to a final volume of 200 µl with 50% ethanol:ultrapure H_2O. Cleaned glass microscope slides were cut into 1×1 cm^2 pieces, and sputter-coated with 50 nm of gold.

Tip functionalization

Standard silicon nitride AFM probes (Digital Instruments, Santa Barbara, CA) were sputter-coated with ~50 nm gold, then incubated in TIP stock solution for 1 h, and rinsed lightly with 50% water:ethanol prior to use.

Microarray construction

The 50 nm gold-coated glass chips were affixed to standard AFM stubs using double-sided tape. Between the tape and the glass a 200 mesh electron microscopy (EM) index grid (Ted Pelco) was placed, to find previously visited areas with ease (27). Pulled glass micropipettes were used to deliver droplets of ssDNA to specific locations of the gold-coated chip. Five microliter droplets of stock solution TIP, MATCH, 1MIS and 2MIS DNA were placed in a 2×2 array, ~100 µm^2 in size. The space between spots was ~30 µm. The microarray was rinsed with 50% ethanol:ultrapure water before use. Immobilization was confirmed with X-ray photoelectron spectroscopy (XPS) (data not shown).

Force spectroscopy

An Extended Multimode Nanoscope IIIa AFM (Digital Instruments, Santa Barbara, CA) was used. The TIP-coated probe and the ODN microarray were placed under phosphate-buffered saline (PBS, pH = 7.4, Sigma) inside the contact mode fluid cell. Evaporation was minimal (relative humidity was 50–60%) but the buffer was occasionally added through the in port of the fluid cell.

Force curves were collected at a rate of 0.1 Hz, over a Z distance of 250 nm. Five hundred curves were collected at a time at each unique spot in the microarray. This was then repeated twice.

The cantilever spring constant was determined using the Cleveland method (20), then used to obtain rupture forces, in concert with cluster analysis. Details of the cluster analysis are given in the Supplementary Material.

RESULTS

The microarray of ODNs and the TIP-coated probe are mounted in the AFM fluid cell containing PBS at pH = 7.4.

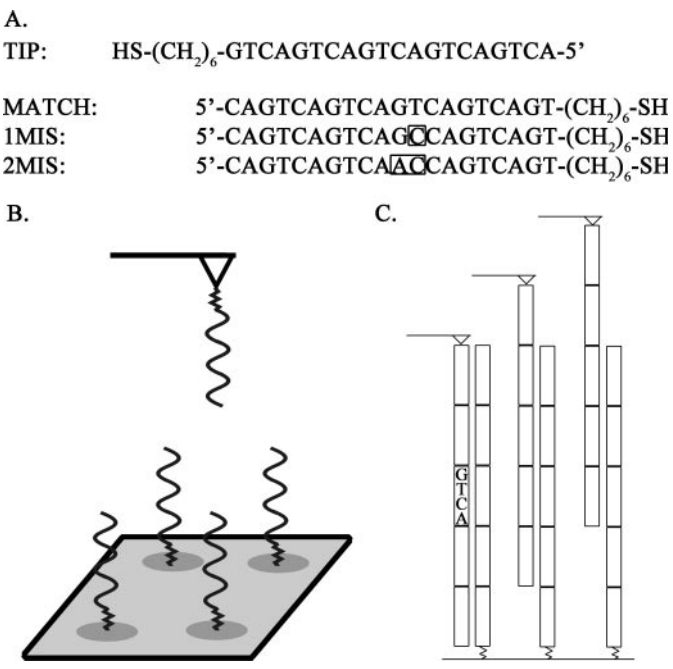

Figure 1. The design. (**A**) The ODNs used have a hexacarbon thiol-terminated linker attached to the 3′ end. (**B**) The TIP sequence is attached to the AFM probe tip. The MATCH, 1MIS, 2MIS and TIP (control) are arrayed on the gold substrate. (**C**) In FS, depiction of three possible overlaps between the TIP and the ODNs: 60, 80 or 100%.

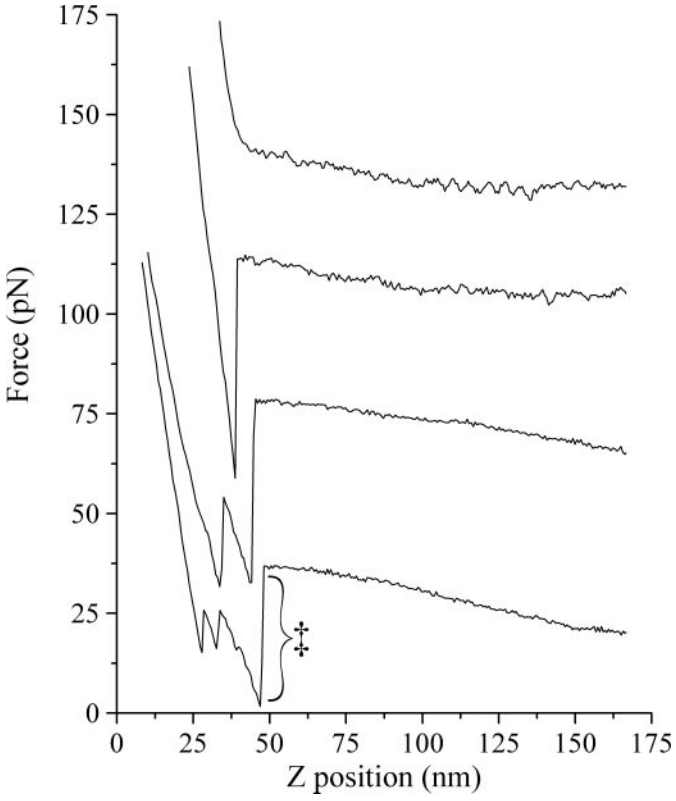

Figure 2. Typical force curves corresponding to none to three rupture events (top to bottom). The final rupture event, due to breaking a single pair interaction, is the only event counted.

Multiple helices can form between TIP and MATCH, 1MIS, and 2MIS (Figure 1). This corresponds to 12, 16 or 20 bp for the case of TIP and MATCH; 11, 15 or 19 bp for TIP and 1MIS; and 10, 14 or 18 bp for TIP and 2MIS. These sets of base pairs are of the same physical length, but the duplexes formed between TIP and 1MIS contain a single base mismatch in the middle of the duplex, and those between TIP and 2MIS have a double mismatch. Computed thermodynamic parameters (1,28,29) indicate that the 8 or 4 bp duplex are not likely to form at room temperature but if they do form, they would be the same for all three pairs.

Data acquisition was performed by ramping the surface first toward and then away from the probe, while monitoring cantilever deflection. A plot of cantilever deflection versus piezo position is called a force curve (Figure 2). Its most important feature for these studies is the jump that occurs during pull-off due to attractive interactions that hold the tip and sample together as they are pulled apart. The tip remains attached to the sample until the cantilever restoring force exceeds the interaction. The size of this discontinuity ('‡' in Figure 2) is extracted using SPMCON95 (C. M. Yip, University of Toronto). The rupture force, which is the force needed to overcome the tip–sample interaction (20), is the product of the cantilever deflection at rupture and the experimentally obtained cantilever spring constant (20).

Since there are many molecules on both the tip and surface, one has to ensure that the force measured represents a single rupture event. There is no direct evidence for this, but there are several factors to consider. The first is the surface coverage: the thiolated ODNs do not form a close-packed monolayer, but rather a homogenous surface coverage of approximately one molecule for every 50 nm^2 (30). Since the diameter of the tip is ~50 nm, the contact area between the surface and tip is ~250 nm^2; thus, only approximately five molecules can interact for any tip–surface contact. Given the distance constraints of the system, it is quite likely that each of the rupture events between different strands will occur at different times during the separation. Thus, we expect to sometimes see multiple rupture events, and we do (Figure 2). For force curves that show multiple interactions, only the last rupture event is considered for analysis. The second is the size of the forces. The AFM can measure forces ranging from picoNewtons to hundreds of nanoNewtons. Since the forces holding the DNA together consist of H-bonding, base stacking and backbone conformation, and we know that the rupture force of H-bonds are approximately few picoNewtons, then we expect to see forces of the order of tens of picoNewtons. While forces of this size are near or below the routine force resolution expected for an AFM, by carefully, mechanically isolating the AFM and ensuring the system remains in complete thermal equilibrium, such high resolution measurements are possible. Note also that these forces can be loading-rate dependent. This could influence the detailed values determined [e.g. Evans and Ritchie (31)].

Experimental conditions were chosen to mimic physiological salt concentration and pH. Apart from biological relevance, these conditions were expected to minimize gold–gold interactions (from the coating on both tip and surface) that could swamp the desired measurement. To confirm this we took 500 force curves between a gold-coated cantilever and a gold-coated surface. In >90% of the curves, no interaction

(hysteresis) was seen, while the remaining ones showed large, randomly distributed rupture events. Additional controls were also performed: Results between a TIP-coated tip and a bare gold surface, and between a TIP-coated tip and a TIP-coated surface are similar to the gold–gold control.

Five hundred consecutive force curves were taken at each spot (MATCH, 1MIS, 2MIS) on the microarray. This process was repeated twice more on each of these spots for a total of 1500 force curves acquired per ODN pair. We found that 35–45% of the force curves produced measurable rupture events. The rupture force was then extracted from each force curve.

The measured forces ranged from 10 to 85 pN. The largest force that holds TIP/MATCH together is at 83.1 pN for the complete 20 bp double helix. For the TIP/1MIS set, this is 81.4 pN, which corresponds to a complete 20 bp helix with one central AC mismatch, and for the TIP/2MIS set, this is 76.6 pN, corresponding to a complete 20 bp helix with a central AC/CA double mismatch.

DISCUSSION

Large data sets were collected for detailed analysis. The sequence chosen was expected to form three unique stable helices (12), and a large data set is expected to define these interactions better, and enable fitting of the histogram of forces to three overlapping Gaussians. What we observed was exactly the opposite: As more data were taken, more peaks appeared in the histograms, indicating that we are seeing breaking of more than the three main helices expected initially. Simulations of the DNA pulling experiment showed that as the end-to-end distance of this complex increased, the end bases stretched out of range for H-bonding and base stacking; yet, the internal bases remained close enough together to maintain all of their interactions (13). Since the strands are capable of completely melting at any time during this process, we postulate that the force spectra contain information not just on the separation of 12, 16 or 20 bp, but of every number of base pair from 20 down to 9. Extending this observation further, we propose that during the pulling event in FS, single or multiple bases at the ends of the double helix lose their double-stranded character before the actual rupture occurs. This means that instead of only observing events corresponding to the breakage of 12, 16 and 20 bp for the MATCH set, we will observe the breakage of every possible number, from 9 to 20 bp (assuming that 8 bp is not stable enough, according to thermodynamic considerations).

Thus, the FS histogram will correspond to as many as 12 different rupture events and analysis by fitting to multiple Gaussians is not reasonable. In addition, histogram analysis itself has considerable bias since small changes in bin size and the starting value for binning can drastically affect the appearance of the histogram. Thus, while histograms can show trends in data, it is difficult to extract precise numbers from them. We resorted to cluster analysis, which enables grouping of data, into homogeneous clusters that are highly heterogeneous when compared to one another. The results are displayed as a dendritic tree, and placing histograms under this tree shows an excellent match between the groups and histogram peaks. In addition no bias has been added to the peak assignment. The statistical validity of the approach is measured by a cophenetic correlation function (c.c.f.). All of our trees are statistically valid, having a c.c.f. of over 0.8. (Details of the cluster analysis are in the Supplementary Material.)

Using a mean Euclidian distance clustering algorithm, we take 13 clusters for TIP/MATCH, 12 clusters for TIP/1MIS and 11 clusters for TIP/2MIS. Taking TIP/MATCH as an example, this accounts for the clusters representing 9–20 bp, and an extra cluster for the larger, non-specific rupture events. Similarly, TIP/1MIS accounts for 9–19 bp plus a cluster for large events, and TIP/2MIS accounts for 9–18 bp plus one extra cluster.

The dendritic tree was computed for each data set. The average value of each tree branch was then taken to represent the branch and used in further analysis. Due to the large size of the data set, the error in this measurement is quite small, ranging from 0.1 to 0.6 pN. The results for all three sequences are summarized in Figure 3, where the rupture force is plotted in two ways, as a function of the expected number of H-bonds in the duplex, which is the same as the number of paired bases in the duplex (Figure 3A), and the number of stacked bases in the duplex (Figure 3B). When assigning the number of stacked bases, we counted the mismatched base pairs as still being stacked, although not H-bonded. For example, a duplex

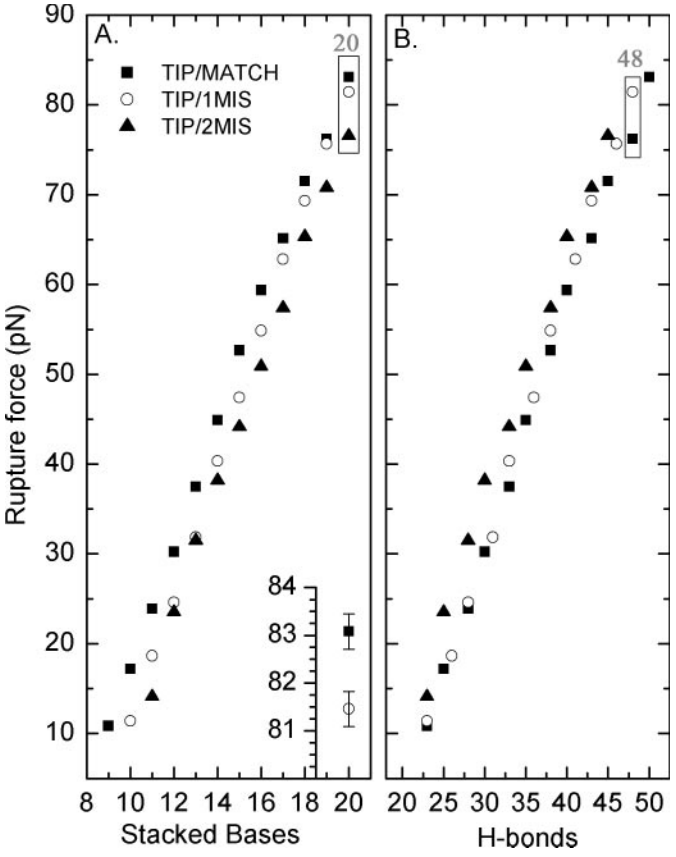

Figure 3. The rupture force for MATCH (square), 1MIS (circle) and 2MIS (triangle) against (**A**) the number of stacked bases in the helix at the time of the rupture event, and (**B**) the number of H-bonds present. Inset of (A), typical error bars. The box labeled '20' shows the rupture force for 20 stacked bases over the three data sets. The box labeled '48' shows the rupture force for 48 H-bonds in the TIP/MATCH and TIP/1MIS data sets.

of 20 aligned bases with one mismatch will be counted as having 19 bases H-bonded, but 20 bases stacked.

The data in Figure 3 is roughly linear, with an *x*-intercept at ~8.0 stacked bases, or 18.0 H-bonds. An 8 bp duplex is predicted to have a rupture force of ~0.02 pN, which is too small to measure. This result is consistent with the thermodynamically based assumption that we should not be able to measure a duplex ≤8 bp.

The data shows that for a given number of stacked bases, the rupture force of TIP/MATCH is greater than that of TIP/1MIS, which in turn is greater than that of TIP/2MIS. This is consistent with our model. Consider the example of 20 stacked bases in a 20 bp double helix (box in Figure 3A). All the duplexes have all their bases stacked, but the H-bonding contributions are different: for TIP/MATCH, all the base pairs are H-bonded, while for TIP/1MIS, the H-bonding contribution due to a single AT base pair is missing. For TIP/2MIS there are two missing H-bonding contributions, due to an AT and a GC pair. Thus, for the same number of total base pairs, the TIP/2MIS is expected to need a smaller rupture force than TIP/1MIS, which in turn is smaller than TIP/MATCH.

On the other hand, for the same number of H-bonds, TIP/2MIS has a greater rupture force than TIP/1MIS, which is greater than TIP/MATCH. For example, in the case of 48 H-bonds (box in Figure 3B), the rupture force of TIP/1MIS is larger than that of TIP/MATCH although they have the same number of H-bonds. This is due to the position of the broken H-bonds. In the TIP/1MIS case, all the 20 bp of the two strands are fully aligned, with two broken H-bonds in the center of the duplex due to the AC mismatch, and the bases on both sides of the mismatch are expected to still interact with each other since they are in a constrained environment. In the TIP/MATCH case of the 48 H-bonds, the broken H-bonds are at the end of the duplex, and the base at the end has less stacking and backbone constraint holding it in place. Similarly, for a fixed number of H-bonds, the rupture force for TIP/2MIS is larger than the corresponding rupture forces for TIP/1MIS and for TIP/MATCH.

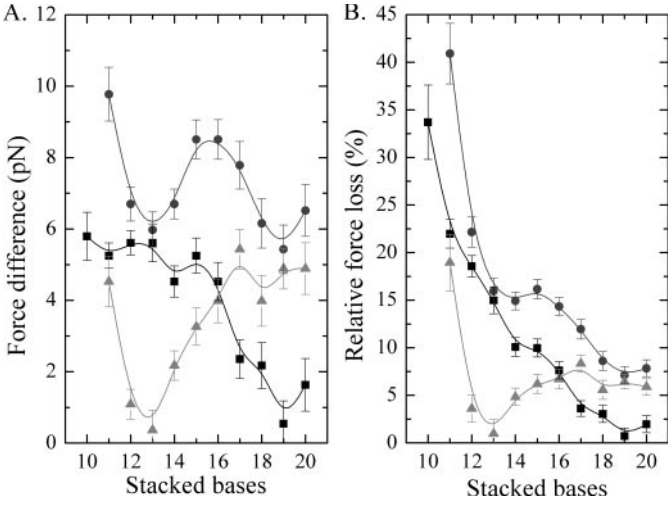

Figure 4. (**A**) Absolute and (**B**) relative force cost of H-bonding for AT (square), AC/TG (circle) and GC (triangle) depending on the helical length.

The magnitude of the forces is in good agreement with previous work on short DNA duplexes ruptured under similar buffer conditions. Strunz *et al*. (16) obtained rupture forces of 25 and 40 pN for DNA duplexes of 10 and 20 bp, respectively. Our values of 20 and 80 pN are within the uncertainties in the spring constants. A 16 bp duplex at 27°C in a saltier buffer (~230 mM) had an unbinding force of ~45 pN (32). Our experiment has the 16mer coming apart at 55 pN. In a low-molarity Tris buffer (20 mM Tris and 10 mM $MgCl_2$) a 12 bp duplex ruptured at ~35 pN (15), but was 25 pN in our study. Furthermore, when experiments were performed on long DNA (i.e. thousands of base pairs) it was found that the double helix is unzippered in a stick/slip type of motion (33). These sticks of 10–20 pN in force occur in >5–10 bp, which correlates very well with our results (33).

Extending these considerations further, we note that the two plots in Figure 3 give information about the two different types of forces that hold the double-stranded DNA together. The plot of rupture force against stacked bases (Figure 3A) provides information about H-bonding contributions, while the plot of rupture force versus H-bonds (Figure 3B) gives information about non-H-bonding interactions, which will be referred to as base-stacking interactions.

H-bonding interaction forces

The difference between data sets with the same number of stacked bases is due to the positional force cost of AC mismatches along the duplex since the bases remain stacked, although the H-bonds are absent (Figure 3A). Three different H-bonding contributions can be determined: (i) TIP/MATCH − TIP/1MIS is the rupture force cost of breaking the H-bonds of an AT base pair; (ii) TIP/MATCH − TIP/2MIS is the cost of breaking AC/TG H-bonds and (iii) TIP/1MIS − TIP/2MIS is the cost of breaking the GC H-bonds, *after* the H-bonds of an AT base pair had already been broken. These differences were calculated and reported in absolute terms (Figure 4A) and as a percentage of the total rupture force (Figure 4B).

The difference (TIP/MATCH − TIP/1MIS) is a measure of the difference between having a matched AT base pair, and an AC mismatched base pair, and is thus the force cost of having an AC mismatch. The duplexes with an AC mismatch have the same number of stacked bases, but two lesser H-bonds and thus, in a simple picture this difference corresponds to the force it takes to break two H-bonds. Comparing TIP/MATCH and TIP/1MIS at different numbers of stacked bases is equivalent to looking at the force cost of having an AC mismatch at different locations along the duplex. As the number of stacked base pair decreases, the position of the AC mismatch moves from the middle toward the end of the duplex. For a 20 bp duplex, with the mismatch near the middle, the force cost is 1.7 pN. A 16 bp duplex, where the mismatch is closer to one end, gives a force cost of 5.0 pN. The shortest measurable duplex, with the mismatch at the very end has a force cost of 5.6 pN. In absolute terms, the force contribution from H-bonds of single AT base pair is small when it is in the middle, and becomes larger as it moves toward the edge of the double helix.

Similarly, the cost of converting AC/TG into AC/CA base pair is determined by calculating the difference in rupture

force between TIP/MATCH and TIP/2MIS. In this case, there is a difference of five H-bonds. For the 20 bp duplex, where the mismatches are near the center of the duplex, we find a force cost of 6.5 pN. As the mismatched pair moves to toward the end, the cost increases, such that for the 16 bp duplex, we get a value of 7.4 pN. For the 12 bp duplex, where the mismatch is at the very end of the duplex, we find an 8.2 pN cost. Thus, the absolute force contribution of an AC/TG pair is 6.5 pN in the middle of a double helix, and increases slightly as the position moves toward the end.

In summary, in longer duplexes the cost of breaking a single AT H-bonds is relatively small, while that of breaking the AC/TG H-bonds is large. However, as we move to shorter helices, the cost of breaking a double helix with just a single AC mismatch becomes comparable to (although still less than) breaking a double helix containing a double mismatch. The inferred cost of breaking a GC after an AT is greater than the AT alone for long helices, but becomes less than the AT for shorter helices.

Strunz *et al*. (16) estimated that the introduction of a single mismatch to a 10 bp duplex pulled at 300 nm/s, will give a difference of 7 pN in the unbinding force, which agrees well with our value of 6 pN at a pulling rate of 50 nm/s.

The proportional cost of the mismatches with regard to breaking H-bonds can be calculated as the rupture force cost per maximum rupture force. The percentage loss in a long helix for a single AC mismatch is low (2–5%), while the cost of the AC/CA mismatch is about twice that (4–10%). As the helix shortens and the mismatch(es) move(s) toward the end, the relative cost of the single AC mismatch increases up to 35% and that of the double mismatch to 42%. The inferred cost of the second mismatch is relatively consistent throughout, with a range of 2–7%. This means that if the base is at or near the end of a double helix, H-bonding plays a greater role in stabilizing the double helix.

Stacking interaction forces

The plot of the number of H-bonds versus rupture force (Figure 3B) can be used to examine the stacking interactions since differences in rupture forces, for the same number of H-bonds, are due to differences in the number of stacked bases. For complexes containing the same number of H-bonds, the TIP/1MIS rupture forces are larger than those of the TIP/MATCH, and the TIP/2MIS rupture forces are larger than both the TIP/1MIS and the TIP/MATCH. These measurements are reported both in absolute and relative terms: Figure 5 shows the cost of unstacking an AT base pair, the cost of unstacking AC/TG base pairs and the inferred cost of unstacking a GC base pair after unstacking an AT base pair. The absolute force cost of unstacking an AT base pair decreases as the number of H-bonds in the double helix decreases, from 5.5 pN at 45 H-bonds to 0.5 pN at 23 H-bonds. The cost of unstacking AC/TG is higher than the cost of unstacking an AT base pair. In general, this force cost increases slightly as the number of total H-bonds decreases, starting at 5 pN for 43 H-bonds and reaching a maximum of 8 pN at 29 H-bonds. As the number of H-bonds decreases further from 29 to 23, the absolute cost of unstacking AC/TG decreases until it reaches 3.2 pN at 23 H-bonds. It follows that the inferred cost of unstacking a GC is low when there are many H-bonds and increases as the size of the helix, and the number of H-bonds, decreases.

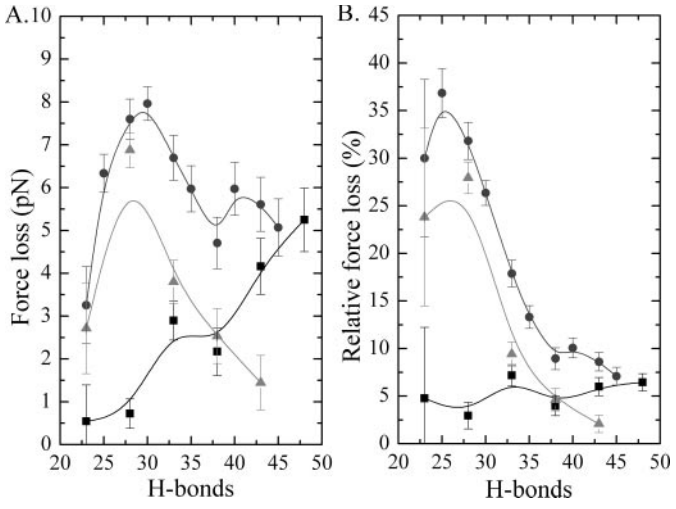

Figure 5. (**A**) Absolute and (**B**) relative force cost of base stacking for AT (square), AC/TG (circle) and GC (triangle), depending on the length of DNA duplex.

The relative cost of unstacking shows a similar trend to the absolute cost. For unstacking an AT base pair, this is ~5–7%, and is independent of the number of H-bonds. On the other hand, the cost of unstacking the AC/TG base pair increases rapidly as the number of H-bonds decreases, from 8% at 45 H-bonds to 38% at 23 H-bonds. This implies that unstacking a single AT base pair does relatively little to destabilize a double helix, while a double mismatch may have substantial effect, depending on helix length. Long duplexes can tolerate the unstacking of 2 bp (AC/TG) without losing much of the force holding them together, but as the number of base pairs decreases, this unstacking dramatically lowers the helix stability with regard to force.

Total force contributions

The difference between points of the same data set (i.e. within TIP/MATCH, TIP/1MIS, or TIP/2MIS) corresponds to the sequential breaking of both the H-bonds and the stacking interactions of the duplex from either end. Since an AT base pair has one less H-bond than a GC base pair, we assume it will break and unstack first. A GC base pair must break next, whether it is beside the broken AT base pair, or at the opposite end of the duplex. Following this convention, we can label each sequential break as an AT or GC base pair. This positional force cost of individual bases is plotted against the length of the double helix and shown in Figure 6. The force costs range from 5 to 9 pN. Within the individual data set, the AT base pair has a rupture force cost of 6.7 pN in the TIP/MATCH set, 6.8 pN in the TIP/1MIS set, and 6.9 pN in the TIP/2MIS set. The GC base pair has a rupture force cost of 6.5 pN in the TIP/MATCH, 7.2 pN in the TIP/1MIS and 6.7 pN in the TIP/2MIS. It is instructive to take the positional average over all three data sets (Figure 7). The force cost of the loss of an AT or a GC base pair ranges from 5.5 to 7.5 pN, with an average of 6.8 pN for AT base pair and 6.7 pN for GC base pair and with no apparent correlation with the length of the

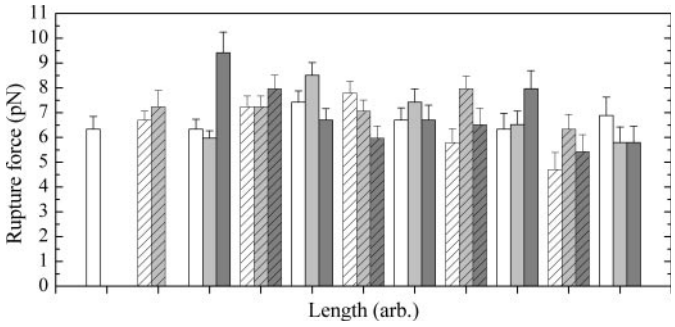

Figure 6. Positional force cost of total loss of base pairing. Difference between consecutive datapoints of the same data set (Figure 4). Solid bars represent AT base pair, hashed bars represent GC base pair, white for TIP/MATCH, gray for TIP/1MIS and black for TIP/2MIS.

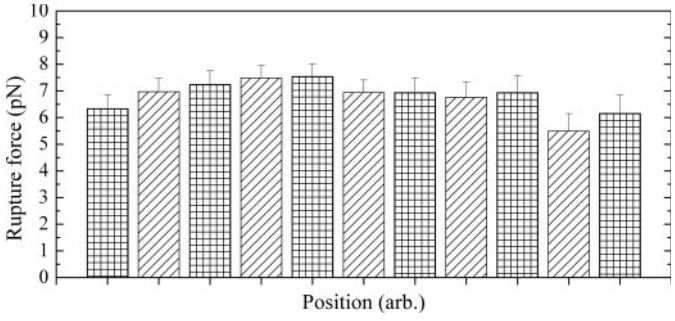

Figure 7. Average cost of base pairing. The average rupture force of AT (meshed) and GC (hashed) base pairs as a function of helical length.

duplex. Since the AT base pair is held together by two H-bonds while the GC base pair is held by three, these numbers would appear to be counter intuitive. However, if one considers that there are both stacking and H-bonding interactions present, the fact that these values are similar (and not in a 2:3 ratio) indicates that the stacking interactions dominate the forces.

Thermodynamic data obtained from melting experiments show that the GC base pair is harder to melt than the AT base pair. This makes sense in the context of the melting experiments, where one varies the temperature and monitors the change in absorbance. The H-bonding part are different for these two types of base pairs, while the stacking interactions are approximately the same; thus, the higher number of H-bonds in the GC base pair will make them harder to melt than the AT base pair. In the force pulling experiment the temperature is held constant, the stacking and torsional constraints on the double helix are manipulated, and the differences in H-bonding could be less important. The force pulling measurement, which is unconventional and does not yield equilibrium parameters, may nevertheless be a more realistic approach to understanding DNA–protein interactions within a cell. Proteins act on DNA not by changing the temperature of the system; rather, they often unwind or pull on double helical DNA when they bind and/or manipulate it.

This notion of the importance of stacking over H-bonding interactions is consistent with the results from the Kool laboratory (34,35). Their studies show that the double helix can form even without H-bonding interactions, hence indicating that base stacking is sufficient to hold the helix together.

CONCLUSION

We presented a novel experimental design to improve the resolution of the AFM for FS. By using a microarray of short ODNs we were able to measure the interactions between different sequences under exactly the same conditions of cantilever, probe tip and solution, hence enabling us to make direct comparisons with the minimal random error. The measured rupture forces are consistent with other recent work, but the force histograms showed a high degree of detail. These details were analyzed using cluster analysis, which allowed resolution of the forces due to rupture events of each base pair along the double helix. By comparing results from sequences that are perfectly matched and those containing mismatches, the force contribution of either an AT or GC base pair was obtained. Surprisingly, these forces were found to be approximately equal, although AT and GC have different numbers of H-bonds. This result indicates that, at least in terms of their contribution to interaction forces in duplex DNA, base stacking is much more important than H-bonding. The approach we developed can be used to further determine the precise force contribution of any base pair in any sequential position. This information will greatly enhance our ability to estimate the forces exerted by proteins on DNA.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank L.-C. Tsui from The Center for Applied Genomics (Toronto, ON, Canada) for the ODNs and C. M. Yip for SPMCON. This work was supported by the NSERC Canada. B.D.S. is grateful for support from the Government of Ontario.

REFERENCES

1. Breslauer,K.J. (1995) Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order–disorder transitions. *Methods Enzymol.*, **259**, 221–242.
2. Breslauer,K.J., Frank,R., Blocker,H. and Marky,L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746–3750.
3. Chalikian,T.V., Volker,J., Plum,G.E. and Breslauer,K.J. (1999) A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques. *Proc. Natl Acad. Sci. USA*, **96**, 7853–7858.
4. Rouzina,I. and Bloomfield,V.A. (1999) Heat capacity effects on the melting of DNA. 1. General aspects. *Biophys. J.*, **77**, 3242–3251.
5. Rouzina,I. and Bloomfield,V.A. (1999) Heat capacity effects on the melting of DNA. 2. Analysis of nearest-neighbor base pair effects. *Biophys. J.*, **77**, 3252–3255.
6. Bockelmann,U., EssevazRoulet,B. and Heslot,F. (1997) Molecular stick-slip motion revealed by opening DNA with picoNewton forces. *Phys. Rev. Lett.*, **79**, 4489–4492.
7. Zhang,Y., Zhou,H.J. and Ou-Yang,Z.C. (2001) Stretching single-stranded DNA: Interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.*, **81**, 1133–1143.
8. Rouzina,I. and Bloomfield,V.A. (2001) Force-induced melting of the DNA double helix 1. Thermodynamic analysis. *Biophys. J.*, **80**, 882–893.
9. Rief,M., Clausen-Schaumann,H. and Gaub,H.E. (1999) Sequence-dependent mechanics of single DNA molecules. *Nature Struct. Biol.*, **6**, 346–349.

10. Clausen-Schaumann,H., Rief,M., Tolksdorf,C. and Gaub,H.E. (2000) Mechanical stability of single DNA molecules. *Biophys. J.*, **78**, 1997–2007.
11. Bustamante,C., Smith,S.B., Liphardt,J. and Smith,D. (2000) Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.*, **10**, 279–285.
12. Lee,G.U., Chrisey,L.A. and Colton,R.J. (1994) Direct measurement of the forces between complementary strands of DNA. *Science*, **266**, 771–773.
13. MacKerell,A.D. and Lee,G.U. (1999) Structure, force, and energy of a double-stranded DNA oligonucleotide under tensile loads. *Eur. Biophys. J.*, **28**, 415–426.
14. Noy,A., Vezenov,D.V., Kayyem,J.F., Meade,T.J. and Lieber,C.M. (1997) Stretching and breaking duplex DNA by chemical force microscopy. *Chem. Biol.*, **4**, 519–527.
15. Pope,L.H., Davies,M.C., Laughton,C.A., Roberts,C.J., Tendler,S.J. and Williams,P.M. (2001) Force-induced melting of a short DNA double helix. *Eur. Biophys. J.*, **30**, 53–62.
16. Strunz,T., Oroszlan,K., Schafer,R. and Guntherodt,H.J. (1999) Dynamic force spectroscopy of single DNA molecules. *Proc. Natl Acad. Sci. USA*, **96**, 11277–11282.
17. Schumakovitch,I., Grange,W., Strunz,T., Bertoncini,P., Guntherodt,H.J. and Hegner,M. (2002) Temperature dependence of unbinding forces between complementary DNA strands. *Biophys. J.*, **82**, 517–521.
18. Boland,T. and Ratner,B.D. (1995) Direct measurement of hydrogen-bonding in DNA nucleotide bases by atomic-force microscopy. *Proc. Natl Acad. Sci. USA*, **92**, 5297–5301.
19. Gibson,C.T., Watson,G.S. and Myhra,S. (1996) Determination of the spring constants of probes for force microscopy/spectroscopy. *Nanotechnology*, **7**, 259–262.
20. Cleveland,J.P., Manne,S., Bocek,D. and Hansma,P.K. (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev. Sci. Instrum.*, **64**, 403–405.
21. Hutter,J.L. and Bechhoefer,J. (1993) Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.*, **64**, 1868–1873.
22. Senden,T.J. and Ducker,W.A. (1994) Experimental-determination of spring constants in atomic-force microscopy. *Langmuir*, **10**, 1003–1004.
23. Butt,H.J. and Jaschke,M. (1995) Calculation of thermal noise in atomic-force microscopy. *Nanotechnology*, **6**, 1–7.
24. Sader,J.E., Larson,I., Mulvaney,P. and White,L.R. (1995) Method for the calibration of atomic-force microscope cantilevers. *Rev. Sci. Instrum.*, **66**, 3789–3798.
25. Sam,M., Boon,E.M., Barton,J.K., Hill,M.G. and Spain,E.M. (2001) Morphology of 15mer duplexes tethered to Au(111) probed using scanning probe microscopy. *Langmuir*, **17**, 5727–5730.
26. Mourougou-Candoni,N., Naud,C. and Thibaudau,F. (2003) Adsorption of thiolated oligonucleotides on gold surfaces: an atomic force microscopy study. *Langmuir*, **19**, 682–686.
27. Markiewicz,P. and Goh,M.C. (1997) Identifying locations on a substrate for the repeated positioning of AFM samples. *Ultramicroscopy*, **68**, 215–221.
28. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A center dot A, C center dot C, G center dot G, and T center dot T mismatches. *Biochemistry*, **38**, 3468–3477.
29. SantaLucia,J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
30. Wang,J. and Bard,A.J. (2001) Monitoring DNA immobilization and hybridization on surfaces by atomic force microscopy force measurements. *Anal. Chem.*, **73**, 2207–2212.
31. Evans,E. and Ritchie,K. (1999) Strength of a weak bond connecting flexible polymer chains. *Biophys. J.*, **76**, 2439–2447.
32. Grange,W., Strunz,T., Schumakovitch,I., Guntherodt,H. and Hegner,M. (2001) Molecular recognition and adhesion of individual DNA strands studied by dynamic force microscopy. *Single Molecules*, **2**, 75–78.
33. Bockelmann,U., Thomen,P., Essevaz-Roulet,B., Viasnoff,V. and Heslot,F. (2002) Unzipping DNA with optical tweezers: high sequence sensitivity and force flips. *Biophys. J.*, **82**, 1537–1553.
34. Delaney,J.C., Henderson,P.T., Helquist,S.A., Morales,J.C., Essigmann,J.M. and Kool,E.T. (2003) High-fidelity *in vivo* replication of DNA base shape mimics without Watson–Crick hydrogen bonds. *Proc. Natl Acad. Sci. USA*, **100**, 4469–4473.
35. Kool,E.T. (2001) Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 1–22.